

IDENTIFYING THE MAINTAINING VARIABLES
OF UNDESIRABLE BEHAVIOR

By

TERESA A. RODGERS

A DISSERTATION SUBMITTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1991

ACKNOWLEDGEMENTS

I would like to thank my committee members Marc Branch, Edward Malagodi, Henry Pennypacker, Carolyn Tucker, and William Wolking for shaping my behavioral repertoire throughout my years in graduate school. I would like to give special thanks to my chair Brian Iwata, for his time and assistance in the preparation of this manuscript and in the development of my career.

Several others have provided assistance in the preparation of this manuscript. Peter Andre, Suneeta Jagtiani, Martina Jonak, Jodi Mazaleski, Richard Smith, Timothy Vollmer and Jennifer Zarccone all served as experimenters during the course of the study. Michael Cripe reviewed earlier drafts of the manuscript. My thanks to each of these individuals.

My husband, Michael Stoutimore, provided much more than editorial comments. His support, guidance, and interactions throughout the years have been invaluable to me.

Finally I would like to thank my father who encouraged me to do my best and who always knew I would make it. This, and all my work, is dedicated to his memory.

TABLE OF CONTENTS

	<u>page</u>
ACKNOWLEDGEMENTS.....	ii
ABSTRACT.....	iv
INTRODUCTION.....	1
GENERAL METHOD.....	27
Subjects.....	27
Human Subjects Protection.....	27
Motivational Assessment Scale (MAS).....	28
Time-referenced assessment.....	29
Behavior-referenced assessment.....	31
Response definitions and direct observation.....	32
Experimenters, observers, MAS raters.....	36
Reliability.....	37
EXPERIMENT 1.....	39
Subjects.....	39
Experimental Design.....	39
Results.....	39
Discussion.....	45
EXPERIMENT 2.....	46
Subjects.....	46
Experimental Design.....	47
Results.....	47
Discussion.....	53
GENERAL DISCUSSION.....	54
REFERENCES.....	60
BIOGRAPHICAL SKETCH.....	63

Abstract of Dissertation Presented to the Graduate
School of the University of Florida in Partial
Fulfillment of the Requirements for the
Degree of Doctor of Philosophy

IDENTIFYING THE MAINTAINING VARIABLES OF UNDESIRABLE
BEHAVIOR

By

Teresa A. Rodgers

August, 1991

Chairman: Brian A. Iwata
Major Department: Psychology

Three methods are use commonly to identify the maintaining variables for aberrant behavior: 1) indirect assessment through interviews, questionnaires, and rating scales; 2) descriptive observational analysis of events as they occur, (e.g., naturalistic observation and field studies); and 3) experimental analysis via exposure to analogue situations.

In the present study, within-subject comparisons were made between the indirect and experimental assessment methods. Experiment 1 compared results from the Motivational Assessment Scale (MAS) to those from a brief (time-referenced) experimental analysis, with 34 developmentally delayed subjects who exhibited self-injurious behavior (SIB). Experiment 2 compared results from the time-referenced

analysis with those from a more extensive behavior-referenced analysis for 10 of the 34 subjects from the first experiment.

The MAS was found to be an invalid assessment tool, in that the results of the two assessments (MAS and time-referenced analysis) matched for only 8 of the 34 subjects. The results of the two experimental analyses matched for 5 of the 8 subjects, indicating that a time-limited experimental analysis provides a better indication of the maintaining variables of SIB than an indirect assessment such as the MAS.

INTRODUCTION

Identification of the maintaining variables for undesirable behavior prior to treatment will result in a more thorough understanding of the behavior and more efficient and effective treatment (Carr, 1977; Demchak & Halle, 1985; Durand & Crimmins, 1988; Iwata, Dorsey, Slifer, Bauman, & Richman, 1982; Johnson & Baumeister, 1978; Sturmey, Carlsen, Crisp & Newton, 1988). An increased interest in and emphasis on assessment of problem behaviors have resulted in a varied technology of assessment and a journal, Behavioral Assessment, devoted solely to the assessment of behavior problems.

In a recent review of the literature, Iwata, Vollmer, and Zarcone (1990) proposed classification of assessment methods based on the type of data collected and the degree of control exerted over behaviors during assessment. The three classifications are: 1) indirect assessment through interviews, questionnaires, and rating scales; 2) uncontrolled observation of naturally occurring events (e.g., naturalistic observation and field studies); and 3) experimental analysis via exposure to test situations.

The most time-efficient method is the indirect assessment. Indirect assessments are based on verbal reports from sources other than the subject. Informants may include family members, teachers, direct-care staff, and counselors/therapists. Chamberlain and Reid (1987) described a checklist called the Parent Daily Report (PDR) and proposed its use as a compromise between other more general checklists and direct observation. They suggested that the problems usually associated with indirect methods of obtaining information, such as systematic bias, requiring the informant to aggregate perceptions over long periods of time (since the problem began, or since working with the individual), and limited utility for assessing short-term changes were eliminated by the PDR. Furthermore, Chamberlain and Reid said that the PDR eliminated these problems by asking the parent over the phone to note whether or not one of 34 behaviors had occurred over the past 24 hours. However, the PDR assessment indicates only the occurrence or non-occurrence of behavior, rather than information with respect to its frequency or environmental correlates.

Durand and Crimmins (1988) argued that assessment methods involving direct observation and manipulation of antecedents and consequences for behavior were too cumbersome and costly in terms of staff time and training. In order to overcome these drawbacks and still identify maintaining variables, they proposed the use of a questionnaire called

the Motivation Assessment Scale (MAS). The sixteen questions on this scale focus on four possible maintaining variables for an individual's SIB: 1) positive reinforcement in the form of attention, 2) positive reinforcement in the form of tangible items (i.e., food, toys, or activities), 3) negative reinforcement (i.e., escape from demands), and 4) sensory reinforcement (i.e., self-stimulation or automatic reinforcement). Durand and Crimmins administered the MAS to a teacher and a teacher's aide for each of 50 students who exhibited SIB. They presented inter-rater and test-retest reliability for the MAS, as well as results from an experiment in which the behavior of subjects in various analogue conditions was compared to the subjects' rating on the MAS. They concluded that the MAS was a reliable and valid indicator of behavioral function, which could be used instead of analogue assessment methods.

A problem with the Durand and Crimmins (1988) study was that the experimenters used only highly trained informants who also were familiar with the subjects, placing a limitation on the use of the MAS similar to that for direct observations; that is, highly trained staff are required. If less well trained individuals or individuals less familiar with the subjects were asked to complete questionnaires, the information provided could be misleading or unreliable. Although Durand and Crimmins reported good reliability, Zarcone, Rodgers, Iwata, Rourke, and Dorsey (in press) failed

to replicate the reliability results of the MAS with both institutional staff and classroom teachers as respondents.

The second method of assessment described by Iwata et al. (1990) is naturalistic observations; these are direct observations of a subject's behavior in the environment where the target behavior typically occurs. The observer may record information in a variety of ways: noting the sequence of events in a behavior log, recording the occurrence or non-occurrence of the target behaviors, or using a behavior code to record the frequency of events. The recording may occur continuously or on a time-sampling schedule. Each of these recording methods has some advantages and limitations. The information log provides a descriptive account of events and their sequence; however, it does not necessarily provide quantifiable data. This method is cumbersome and allows for the recording of only a few events at one time. The occurrence/non-occurrence method is more precise and provides some estimate for frequency, as well as a time-based account of events; however, the estimate of frequency is subject to distortion depending on the duration and rate of the behavior. The behavior code groups responses into classes, and so allows for the recording of several events or responses of several individuals at the same time. However, the grouping is probably based on topography rather than function and may restrict interpretation if the responses or events are functionally different.

There are at least two problems with naturalistic observations as a whole: one is that the controlling variable(s) may not be readily apparent in the environment; another is that the method does not demonstrate functional relations (i.e., the information is correlational). The first problem is difficult to detect and to overcome because the conditioning that resulted in the initial control, either antecedent or consequent, probably occurred at an earlier time. However, some pairing of stimuli and response must continue to occur at least occasionally, or control by the variable would have been extinguished; the variable and its effects might be intermittently observed in the natural environment. For example, it may be that the subject initially emitted some aberrant response and received a great deal of attention from staff. As this response was new for the subject, almost every response may have occasioned more concern from staff. As time passed and staff habituated to the response, they might have provided attention intermittently. Finally, the behavior would have been maintained on a lean schedule of reinforcement. Natural observations would reveal, in this case, that staff rarely attend to the response. However, it is this limited amount of attention that maintains the response. This type of control would be difficult to detect during naturalistic observations. To address the second limitation of naturalistic observations, the lack of information with

respect to functional relations, there is no solution other than to pursue an experimental analysis of the behavior.

Bijou, Peterson, and Ault (1968) described a system that differed from previous descriptive studies in that frequency data were collected rather than data in the form of rating scales, running descriptions of situations, or some other indirect, subjective measurement. The collection of frequency data in the natural environment allowed for the comparison of data from naturalistic and experimental studies. Bijou et al. specified the behavioral and stimulus events to be observed, and recorded data for these events in two ways. One method consisted of logging the occurrence of responses, and sometimes the duration of each response, and the other consisted of registering the frequencies of occurrences and non-occurrences within time intervals.

The observation system outlined by Bijou et al. (1968) eliminated problems of inadequate recording of behavior and limited response definitions, problems common to descriptive or field studies (Johnson & Bolstad, 1973). The data collected are quantified, the responses are observable and measurable, and the methodology produces replicable results. The limitation remains, because only naturalistic observation occurred, that the maintaining variable may not be readily apparent in the current environment.

Another descriptive method used to identify variables controlling aberrant behavior is the scatterplot analysis

developed by Touchette, MacDonald, and Langer (1985). Frequencies for some interval of time were recorded on a scatterplot grid. Different symbols were used for no occurrence of the response during the interval, for low rates of the response, and for high rates of the response. The grid sheet formed a visual display of the data without formal calculations, and time intervals containing high rates of behaviors were identified by the grouping of high rate symbols within and across days. The criteria for low and high rates were determined individually based on the frequency of the response. Activities, staff, and other situational variables were identified for the high rate periods.

Touchette et al. (1985) described the scatterplot as identifying patterns of responding occurring in the natural setting, and these patterns of responding suggested sources of environmental control. Unlike a line-graph display of data that reveals only average rate and necessarily condenses the pattern of the response into larger units, Touchette et al. claimed that the scatterplot keeps the pattern of responding intact. They proposed that the scatterplot would be useful when the target behavior is frequent and there is no obvious correspondence between environmental conditions.

A limitation of the scatterplot method is that the variables identified are molar. The association of a time of day and activity does not provide information with respect to

the specific stimuli or contingencies operating on the behavior. Touchette's suggested treatment, based on a scatterplot analysis, is to rearrange the subject's schedule so that the condition associated with high rates either no longer occurs, or occurs at a lower frequency or different time. This treatment does not take into account the specific aspects of the identified problem situations that maintain the behavior. Re-arranging the schedule may result in an initial decrease in the response; however, when the subject has been in the new situation and is exposed to similar stimuli or contingencies, the problem behavior may reoccur.

Another problem with the scatterplot method is that the grouping of frequencies may distort the response pattern. The actual frequency of response is lost, and if the condensation criteria are inappropriate, important periods of time may be excluded from consideration as problem times. It seems that it would involve little effort to record the actual frequency for each interval, and then determine the high rate and low rate times.

The third approach to assessment described by Iwata et al. (1990) is the experimental analysis. It is the most thorough method for obtaining information about maintaining variables, and involves arranging consequences to test for variables reinforcing the behavior of interest. Two procedures have been called experimental analyses: one is the implementation of several treatments and identification

of the maintaining variable based on the treatment that was the most effective, the other is to provide special arrangements of consequences for the behavior in controlled situations that recreate the possible reinforcement available in the typical environment. Special arrangements of consequences are standardized by the reinforcer classifications (i.e., positive or negative reinforcement); additionally, arrangements of particular stimuli can be individualized based on the subject's specific history.

The first type of test is a post-hoc analysis that may not identify the maintaining variable; the behavior may simply come under the control of the new arrangement of contingencies. An example of such an analysis is the study by Repp, Felce, and Barton (1988), who tested hypotheses regarding the maintaining variables predicted from baseline observations by exposing subjects to two of three different treatments: either extinction of positively reinforced behavior, extinction of negatively reinforced behavior, or increased stimulation. Based on effects of the treatments, they concluded that their hypotheses had been correct for all three of their subjects. Repp et al. (1988) did not provide the baseline data, thereby preventing independent confirmation of the hypotheses. A difficulty of this approach is insuring that adequate data have been collected during baseline to formulate hypotheses. Repp et al. included a naturalistic assessment component to their

analysis; they collected data specific to the possible maintaining variables for each subject. That is, to determine if the behavior might be positively reinforced, they collected data on consequences for behavior including teacher proximity, verbal attention from staff, and touching. For negatively reinforced behavior, they noted what was removed following the behavior. For self-stimulation as the maintaining variable, they noted whether the subject had environmental stimulation available prior to occurrence of the behavior.

Several examples of the second type, (i.e., arranging analogue conditions to test for maintaining variables), have been reported. Carr, Newsom, and Binkoff (1980), after observing subjects in their typical environment, hypothesized that aggressive responses were motivated by escape from demands. They then systematically manipulated environmental variables to test for escape-maintained aggression. The variables manipulated were demands, reinforcement for correct responses, stimuli associated with the end of sessions (safety signals), and escape from the session. Sessions were 5-min long for one subject and 10-min long for the other. Frequency data were collected in 1-min intervals. A reversal design was employed with different demands for both subjects. In addition, the sessions ended with a safety signal.

They found that demands were associated with increased aggressive responses and safety signals were associated with

a decrease in aggressive responses for both subjects. A reinforcement contingency for correct responses resulted in a decrease in aggressive responses. An escape contingency for a targeted non-aggressive response resulted in a decrease in aggressive responses. An increase in aggressive responses occurred with contingent escape from the demand situation. Extinction through removal of the escape contingency resulted in a decrease in aggressive responses. They concluded that the aggressive responses of both subjects were negatively reinforced by escape from demand situations, and they designed treatment conditions based on their results.

Iwata et al. (1982) also attempted to identify functional relations between self-injury and specific environmental events through brief, repeated exposures to a series of standard conditions. The environmental variables manipulated in the conditions were availability of play materials, attention, and demands. Nine subjects were exposed to 4 experimental conditions in a multielement design. An observer recorded the occurrence and nonoccurrence of self-injurious behavior for continuous, 10-s intervals. The dependent variable was the percentage of intervals during which one or more self-injurious responses were scored.

The four conditions were social disapproval, academic demand, alone, and unstructured play. The social disapproval condition was arranged to test for positive reinforcement by

staff attention to self-injurious behavior. The experimenter did not interrupt the SIB, although physical contact was paired with the social disapproval.

The academic demand condition was arranged to test for negative reinforcement by removal of demand situations when SIB occurred. During this condition, educational activities were presented to the subject every 30 s. Contingent upon SIB, the experimenter turned away from the subject for 30 s and the trial was terminated. A change-over delay of 30 s was implemented for each repeated SIB.

The alone condition tested the effects of an impoverished environment. The subject was left alone in a room with the observer behind a one-way mirror in another room.

The unstructured play condition was arranged as a control for the presence of the experimenter, lack of attention for SIB, absence of demands, and the availability of potentially stimulating materials. The contingencies arranged in this condition were the opposite of the demand and social disapproval conditions.

Iwata et al. (1982) were able to identify functional relations between specific environmental conditions and SIB for 6 of the 9 subjects. The functional relations were evinced by the differential patterns of responding. Three of the subjects exhibited no differential pattern of responding across the conditions. Iwata et al. hypothesized that for

these subjects the conditions were not discriminable, the behavior was a function of variables not controlled by the study, or that the behavior was under the control of multiple variables (attention, escape from demands, and stimulation in barren environments).

In contrast to the Carr et al. (1980) study, Iwata et al. (1982) did not base their analogue conditions on observations in the subjects' environment; they developed conditions to test for several possible sources of reinforcement. An advantage of the standard conditions was that they allowed the experimenters to test for more than one controlling variable, whereas the specifically designed conditions of Carr et al. allowed testing of only one controlling variable. A disadvantage is that observations from the natural environment (as in Carr et al.) may have provided more information with respect to the behavior of subjects for whom there was no discernable pattern of responding in the analogue conditions.

Mace and Knight (1986) provided another example of an experimental analysis involving special arrangements of consequences for the behavior. They performed a functional analysis for the pica exhibited by a profoundly retarded male. The variables manipulated were staff interaction and the subject wearing a helmet with face shield. The study was performed in two settings: the assessment occurred in a prevocational classroom with 5 other clients, and the

treatment intervention was implemented in a day area with 8 other clients. Sessions were 15 minutes long, and a 10-s, partial-interval system was employed to collect data.

The assessment baseline contingency was to follow each pica response or attempted response immediately with a mild reprimand and removal of the object from the subject's mouth. The effects of varying rates of interaction were assessed in conditions when objects were removed from the subject's mouth without a reprimand. The amount of social interaction was manipulated in four conditions: 1) noncontingent interaction, the experimenter delivered a prompt to complete the task according to a variable time (VT) 8-min schedule, 2) frequent interaction with continuous eye contact and task instructions on a VT 15-s schedule, 3) limited interaction in which the experimenter faced away from the subject and interacted with respect to the task only on a FT 3-min schedule, and 4) no interaction in which the experimenter faced away from the subject at a different table. The effects of the helmet were analyzed in three separate conditions, all within the limited-interaction contingency: helmet and face shield on, helmet only, and no helmet. The percentage of intervals with pica was related to the rate of interaction; low rates of interaction produced high percentage of intervals with pica. In addition, the percentage of intervals with pica was differentially affected by the amount of protective equipment; no helmet resulted in

the fewest percent intervals of pica. Finally, the experimenters designed a treatment consisting of limited therapist interaction with the subject without a helmet. The treatment reduced the percentage of intervals with pica.

The study demonstrates an interesting parametric analysis of possible antecedent and consequent controlling variables. The subject's behavior systematically changed with the change in contingencies, indicating a functional relationship between the variables manipulated and pica. The rate of the subject's behavior was lowest in the frequent interaction condition and highest in the no interaction condition; the helmet with the faceshield produced the highest rate of pica within the helmet manipulations. No explanation was provided for including limited interaction rather than frequent interaction in the treatment. Mace and Knight combined aspects of the naturalistic observation method and collected data in the subject's home environment. These corroborating data from the natural environment provide support for the conclusions from the environmental manipulations. They found low rates of pica when there was interaction with the subject without the helmet. They further found that staff interacted less frequently with the subject when he wore the helmet and face shield.

Sturmey, Carlsen, Crisp, and Newton (1988) revised and extended the functional analysis of Iwata et al. (1982). They analyzed stereotypic behaviors such as chest patting,

rocking, head rolling, and finger manipulations. The contingencies implemented in each of the four conditions were the same as in Iwata et al.: alone, unstructured play, demand, and social disapproval. A multielement design was employed. For 2 subjects, they found that the rate of stereotypy was highest in the alone condition, and low in conditions in which any distractions were available (i.e., demand, attention, and play). The third subject exhibited essentially undifferentiated rates of stereotypy in all conditions.

The analysis performed by Sturmey et al. (1988) differed from Iwata et al. (1982) in several ways. First, they used momentary time sampling to record occurrence or nonoccurrence every 15 s. They proposed that this measure would give an unbiased estimate of the proportion of time spent behaving. Second, Sturmey et al. had 10-min sessions rather than the 15-min sessions of Iwata et al. This decreased time period may have reduced the subjects' exposure to the experimental contingencies enough to cause the undifferentiated effects obtained by Sturmey et al. The final difference was that only one experimenter conducted all of the sessions in each condition. This may have decreased the discriminability of the conditions and resulted in less differentiation for each subject across conditions.

Durand and Crimmins (1988) included an analogue assessment in their study on the MAS. The assessment

consisted of five conditions: baseline, attention, escape, tangible, and unstructured. Subjects were exposed to each condition for three, 10-min sessions. The baseline condition consisted of requiring the subjects to perform an "easy" task (matching sample items from an intelligence test, the Leiter Scale). The task was determined as easy because only items that the subject was able to match with 100% accuracy were included. Verbal praise contingent upon correct responses and attention in the form of either a demand, praise, or some other neutral comment were provided every 10 s. In addition, contingent on correct responses, tangible items were provided on a variable ratio (VR) 3 schedule. There were no consequences for SIB or other inappropriate behaviors. This was a control condition; it was hypothesized that no aversive stimuli were present and that socially motivated SIB would be low due to the high density of attention.

The attention condition involved the same contingencies as the baseline condition with the exception that instead of providing attention for 100% of the intervals, attention was provided for only 33% of the intervals. This condition was designed to test adult attention as a maintaining variable for SIB.

The escape condition involved the same contingencies as the baseline condition except that the task was more difficult; the subject could respond correctly to only 33% of

the items. This condition was designed to test for negative reinforcement as the maintaining variable for SIB.

The tangible condition involved providing each subject's favorite tangibles every ninth correct answer, while the items were visible constantly. This condition was designed to test for tangible reinforcement as the maintaining variable for SIB.

The unstructured condition involved placing the subjects' favorite foods and toys within reach, having an adult interacting with the subject, and having work materials available for the subject. This condition was designed to test for sensory consequences as the maintaining variable for SIB.

Durand and Crimmins (1988) found that each of the 8 subjects exhibited a high rate of responding in one of the experimental conditions. Two of the subjects had higher rates of responding in the unstructured condition, and two had high rates in the attention condition. Two other subjects had higher rates in the escape condition, and two had high rates in the tangible condition.

There were significant problems with the experimental conditions in the Durand and Crimmins (1988) study, as well as with the conclusions drawn from them. Durand and Crimmins described their conditions as designed to produce increases in SIB, if SIB was maintained by the variable manipulated in each condition. However, none of the manipulations were made

contingent upon the SIB. That is, in the attention condition, attention was not delivered contingently, as if it were the reinforcing stimulus maintaining the SIB. Instead, attention was delivered every 30 s, regardless of the subject's behavior. In such an arrangement one might expect to reinforce SIB accidentally; however, there would be no arrangement that would determine if the SIB was sensitive to attention as a reinforcer. The reinforcers for the other conditions were similarly arranged on a fixed-time schedule without a response requirement. Therefore, instead of the high rates of SIB that Durand and Crimmins predicted, one might argue that the frequency of SIB would be depressed because reinforcers were delivered regularly without the SIB requirement. Because the stimuli were not delivered contingent on SIB, they were not demonstrated to be reinforcers or maintaining variables for SIB. For all of the conditions there are alternate explanations for the results obtained. The Durand and Crimmins unstructured condition may have resulted in high rates of SIB because the people in the room served as discriminative stimuli either that SIB would be reinforced with escape or by attention from staff. In contrast, the Iwata et al (1982) "alone" condition placed the subject in a room without any play objects and with no opportunity for social interaction, eliminating escape and social reinforcement as controlling variables and strengthening an argument for sensory stimulation. The two

conditions, therefore, presented almost opposite environmental contingencies, and a comparison of responding during each condition would provide additional information regarding sensory maintained SIB.

Steege, Wacker, Berg, Cigrand, and Cooper (1989) assessed potential reinforcers and variables maintaining SIB for 2 subjects. Their study was an extension of Iwata et al. (1982) in that a stimulus preference was included in the assessment procedure, and the reinforcers identified were then arranged contingent upon adaptive behavior during treatment.

It was determined by interviews with the classroom teacher that SIB (handmouthing and biting) occurred most frequently when the first subject was alone. The assessment of maintaining variables included three conditions. In two of the conditions there was no contingency for SIB and the subject was left alone except for the observer. The third condition was a combination of the Iwata et al. demand and attention conditions. Data were recorded during 6-s intervals, and sessions lasted 15 min. Steege et al. (1989) considered the first three sessions in each of the conditions as the assessment. SIB did not occur during the demand/attention condition, and occurred at approximately equal rates in both of the alone conditions. Treatment involved differentially reinforcing an alternative (DRA) response (pressing a microswitch), and low rates of SIB were

maintained across all conditions. Steege et al. concluded that SIB was maintained as a self-stimulatory behavior.

The assessment for the second subject more closely approximated the Iwata et al. (1982) assessment. Three of the four conditions--alone, demand, unstructured play--were included, and a response cost condition was added. The response cost condition was said to approximate conditions in the natural environment that were reported by the parents to result in high rates of SIB. Toys were available in the room. If the subject played with a toy for 20 s, the experimenter removed the toy and scolded the subject for playing with it. Steege et al. (1989) arranged for the subject to be exposed to these conditions briefly (4, 10-min sessions of each condition). None of the data from the assessment were presented in the study, although Steege et al. reported that SIB occurred exclusively during the demand condition. They concluded that the subject's SIB was maintained by negative reinforcement and designed a treatment in which SIB resulted in the subject being redirected to a task, and correct responses produced an identified reinforcer. The SIB was reduced within the first three sessions after treatment was instituted.

Wacker et al. (1990) also provided brief exposure to contingencies within a functional analysis framework. They exposed subjects for a total of 90 min or less to conditions of the Iwata et al. (1982) functional analysis. The

maintaining variables were identified as positive reinforcement for 1 subject, automatic reinforcement for another subject, and negative reinforcement for the third subject. Although it appeared that the responding of subjects was under the control of the contingencies arranged in the various conditions, the brief exposure raises the possibility of a transient trend that would have changed had the subject been exposed to the conditions for longer periods of time.

Steege et al. (1990) also implemented brief functional analyses to determine that the SIB of 2 subjects was maintained by negative reinforcement. Subjects were exposed briefly (7, 10-min sessions and 4, 10-min sessions) to each of the four Iwata et al. conditions. They noted that several sessions were terminated early and so lasted less than 10 min.

Even though analogue assessments allow the identification of functional relations, it might be possible to argue that these prosthetic assessment situations provided "training" of SIB in the demand or attention conditions. If the subject is responsive to the contingencies provided, attention or escape from a task, then presenting stimuli contingent on SIB will increase the probability of future SIB responding. This will occur even if other variables are responsible for the occurrence of the SIB in the natural environment, because in the assessment situation all other

variables are controlled, and the reinforcers available are contingent on SIB. However, if the subject responds quickly to the contingencies, and reports from the subject's normal environment support that the SIB serves an escape function, it is unlikely that the assessment develops "new" reinforcement contingencies for the SIB. In addition, if the behavior were to come under the control of new reinforcement contingencies, one might expect the rate of SIB to increase over the course of the assessment forming an acquisition curve. This acquisition curve would serve as an indicator that the SIB had been "trained" rather than a maintaining variable identified.

In summary, there are three general methods of assessment: indirect information gathering, observation in the natural environment, and experimental manipulation of possible controlling variables. It is often assumed that data obtained from the different assessment techniques are correlated, although experimental support for this assumption is generally lacking. However, one comparison of an indirect method of assessment and direct measurement of subjects' behavior was conducted by Durand and Crimmins (1988). They assessed the validity of the MAS for identifying the variables controlling SIB for 8 subjects selected from an original sample of 50 subjects. For each subject, a single motivating condition was identified by the MAS. The results of the questionnaire were compared to those obtained from a

functional analysis was similar to but briefer than that arranged by Iwata et al. (1982). Durand & Crimmins (1988) concluded that the MAS provided a valid prediction of results of the analogue conditions; however, a close examination of their data indicates that of the eight subjects in the experiment, the results for four were inconclusive. One subject had less than 15% intervals scored in the highest condition, and another subject had little difference between the highest category and the next highest category in the percent intervals scored. Two of the subjects only exhibited SIB during the demand condition with low percent intervals scored.

Several of the analogue studies described (Sturmey et al., 1988; Durand & Crimmins, 1988; and Steege et al., 1989) exposed subjects to the assessment conditions for either very few sessions, very brief sessions, or both few and brief sessions. The assessments were time-referenced rather than following a criterion of stability of responding (Sidman, 1960). The validity of these brief-exposure assessments has not been determined and could be assessed through a comparison to behavior-referenced assessment.

The present study compares three assessment techniques: an indirect method, as in the Durand & Crimmins (1988) study; and two analogue analyses, a time-referenced assessment and a behavior-referenced assessment. The results of the MAS and the two functional analyses are compared in order to

determine the validity of the MAS as a predictor of results from the analogue analyses. A more important comparison is the results of a time-referenced exposure to functional analysis conditions and a behavior-referenced exposure, allowing determination of the validity of the time-referenced assessment. The specific questions asked in this study are: 1) Is the MAS a valid and useful tool for determining the maintaining variables of a subject's SIB?; and 2) Is a time-referenced exposure to the analogue conditions a valid and useful indicator of the maintaining variables of a subject's SIB? Two experiments were performed in order to answer these questions. Experiment 1 compared the results of the MAS to the results of a time-referenced analogue assessment. Because of the numerous procedural problems previously specified and, therefore, the limited possible inferences from the Durand and Crimmins (1988) study, the conditions of the time-referenced assessment had contingencies arranged as in the Iwata et al. (1982) study. However, the MAS is compared to a time-referenced assessment, rather than a behavior-referenced assessment, to provide a more direct replication of the Durand and Crimmins study. In addition, the time-referenced assessment arrangement is a combination of all previously described limited assessments (Steege et al. 1989; Steege et al. 1990; Sturmey et al. 1988; Wacker et al. 1990). The comparison between the MAS and a behavior-referenced assessment will determine if the MAS is a valid

indicator of maintaining variables as identified in a behavior-referenced assessment. Experiment 2 compared the results of a time-referenced to a behavior-referenced assessment in order to determine if the time-referenced procedure is a valid assessment methodology.

General Method

Subjects

Subjects were drawn from a pool of residents of a state institution for the mentally retarded who had been referred for assessment and/or treatment of SIB. All subjects were reported to be functioning in the severe to profound range of mental retardation and had limited communication and self-care repertoires.

Human Subjects Protection

In order to assess the differential effects of the environment on self-injury, the study required that subjects be allowed to engage in minor self-injurious behavior for brief periods of time. All of the procedures were reviewed and approved by an institutional review board, as well as the institutional behavioral program review committee and medical staff. The following safeguards were employed to reduce the risk of physical injury as a result of the SIB exhibited during assessment. First, each subject received a complete medical examination by a physician, which assessed current physical status, ruled out organic causes for the SIB, and established a session termination criterion based on the severity of the behavior. Physicians occasionally observed sessions in order to assess a subject's SIB as it occurred. If a subject's physical condition or level of responding met the criterion for session termination, the session was stopped and the SIB was prevented via physical or mechanical

restraint. Further sessions were postponed if the SIB resulted in an open wound. During the course of Experiment 2, session termination and/or suspension only occurred one time during the time-referenced assessment for 2 subjects, and two times with 1 subject during the behavior-referenced assessments.

Assessment Conditions

Motivational Assessment Scale (MAS). The MAS (Durand & Crimmins, 1988) consists of 16 questions about events surrounding self-injurious behavior. Durand and Crimmins identified four categories of reinforcement: 1) positive reinforcement in the form of adult attention, 2) positive reinforcement in the form of a tangible item, 3) escape from demand situations, and 4) sensory reinforcement. Each category is represented by four questions on the MAS. The answers, indicating how often SIB occurs under the circumstances described, are specified on a six-point, Likert-type scale where (0) = never, (1) = almost never, (2) = seldom, (3) = half the time, (4) = usually, (5) = almost always, and (6) = always. The numbers associated with each answer for each of the four questions related to a category were then added, and an average rating for each category was computed. The category with the highest average indicated the source of reinforcement for the SIB.

The MAS was administered as described in Durand & Crimmins (1988) with the following exceptions. The staff

responders in the present study consisted of cottage supervisors and direct-care staff, instead of classroom teachers. To avoid problems such as timely return of the questionnaire and possible reading difficulties of staff, the questions were read to each staff person by a graduate research assistant. The MAS was administered in as private a setting as possible, and the only prompts for responders were to repeat the questions and to repeat the possible answers. The staff person answering the questions had all of the possible answers in front of him/her on a typed card, and the administrator of the questionnaire read the question and asked the staff to provide one of the possible answers, ranging from zero ("never") to six ("always"). When a staff person had difficulty answering a question, the research assistant repeated the question and asked the staff person to give the best answer. If the staff person's reply did not coincide with one of the six possible answers, the administrator asked which of the six answers came closest to his/her response.

Time-referenced assessment. Three sets of contingencies (demand, attention, and alone) were in effect during the time-referenced assessment. A session consisted of exposure to the contingencies associated with one of the conditions. Each session lasted 5 min, and each condition was implemented three times. Sessions were sequenced in a random order with the restriction that one condition could not be repeated

until all three conditions had been presented. The time-referenced assessment was completed in a private room on the subject's home cottage.

In the demand condition, the therapist, subject, and observer(s) were present. Academic or self-care materials were present in the room. Every 30 s the therapist presented an instruction or task to the subject. Contingent upon SIB, the instruction or task was terminated for 30 s. Praise was delivered contingent on compliance with the instruction. All other behaviors were ignored. This condition was designed to allow escape from the demand contingent upon SIB; if escape functioned as a reinforcer, the rate of SIB was expected to be higher in this condition than in other conditions.

In the attention condition, the subject, therapist, and observer(s) were present. No toys, work materials, or other potentially stimulating items were present. Contingent upon SIB, the therapist verbally attended to the subject, providing statements of social disapproval or concern. In addition, attention consisted of response blocking for 1 s or otherwise touching the subject. All other behaviors of the subject, appropriate or inappropriate, were ignored. This condition was designed so that the only contingency was attention for SIB. If attention functioned as a reinforcer, the rate of SIB would be higher in this condition than in the other conditions.

In the alone condition, the subject was alone in the experimental room with no materials available. Although at least one observer was present, the observer did not interact with the subject or with any other observer who might have been present. No eye contact was made between the subject and the observer. If the subject approached the observer, the observer moved away without obviously reacting in any other way to the approach of the subject and attempted to maintain a distance of at least 3 feet. In this condition there were no contingencies for SIB or any other behavior, and there were no distracting or stimulating items available. It was expected that if SIB was automatically reinforced, the rate would be highest in this condition.

Behavior-referenced assessment. The behavior-referenced assessment was comprised of four conditions: alone, demand, attention, and play. The behavior-referenced assessment was completed in a cottage established for the treatment of residents who engaged in SIB. There were never more than two other subjects present on the cottage at any one time, and subjects spent less than two hours per day on the cottage. Session length was 15 min across all conditions. At least two sessions and as many as four were run each day. Sessions were sequenced in a random order with the restriction that one condition could not be repeated until all four conditions had been presented.

The contingencies implemented for the demand, attention and alone conditions were the same as described for the time-referenced assessment. In the play condition, various toys were present in the room with the subject and observer(s). A therapist approached the subject and either presented a toy or interacted with the subject on a 30-s, fixed-time schedule. All SIB was ignored and if a self-injurious response occurred within 5 s of the end of a fixed-time interval, the therapist delayed the delivery of attention for 5 s to avoid reinforcing the SIB. This condition was designed to control for the effects of having staff and materials in the room, and to provide consequences for appropriate behavior. SIB was expected to have the lowest rate in this condition.

Figure 1 presents the conditions and controlling variables that correspond for each of the three assessments. For example, the MAS conditions of Attention and Tangible correspond to the Attention condition for both the time-referenced and behavior-referenced assessments, the controlling variable for these corresponding conditions is environmentally located positive reinforcement.

Measurement

Response definitions and direct observation. Table 1 contains a listing of the specific self-injurious responses observed for each subject, along with the operational definitions used in collecting data.

Figure 1. The assessment conditions and controlling variables that correspond for each of the three assessments are presented. The arrows connect corresponding conditions. Only the behavior-referenced assessment had a control condition.

CONTROLLING VARIABLE

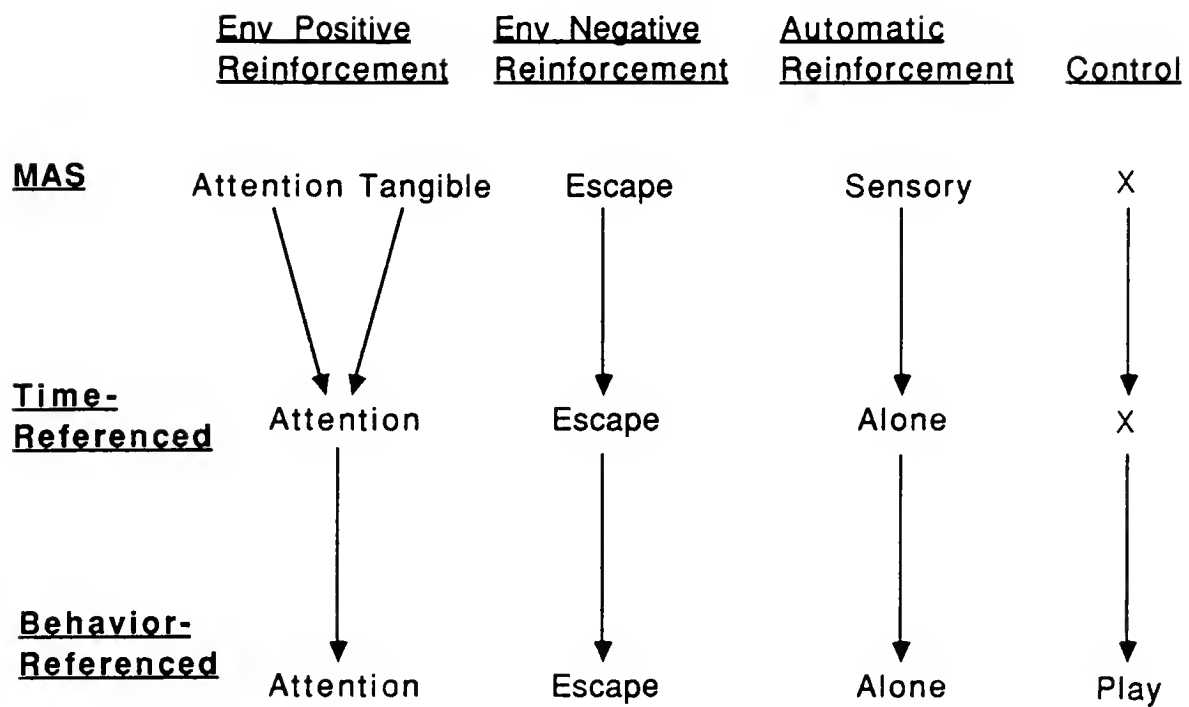


Table 1

Observer definitions of subject's self-injury

<u>Response</u>	<u>Definition</u>	<u>Number of Subjects</u>
Arm/hand banging	Forceful contact of the arm/hand with a stationary object	5
Face slapping	Forceful contact of the open hand with the face	25
Hair pulling	Closure of the fingers and thumb on hair with a pulling motion away from the head	5
Hand mouthing	Insertion of one or more fingers into mouth, or tongue touching any part of hand	6
Head banging	Forceful contact of the head with a stationary object	18
Self-biting	Closure of the upper and lower teeth on the flesh of any portion of the body	11
Self-kicking	Forceful contact of a foot to any portion of the body	5
Self-scratching	Gouging movement of fingernails across skin, leaving red mark or wound	5
Throat-poking	Forceful jab of throat with finger or hand	2

During each session, an observer recorded occurrences of SIB, compliance, disruption, and aggression in continuous 10-s intervals. Occurrences of significant experimenter responses (e.g., instruction or attention) were recorded as well. Observers recorded responses on a hand-held computer (Assistant Model AST102). The dependent variable of interest was the number of responses per minute for each session. The mean number of responses per minute was then computed for each condition of the time-referenced and behavior-referenced assessment.

Experimenters, observers, MAS raters. Experimenters (i.e., therapists) were eight graduate research assistants. Each experimenter was provided with a written description of the assessment conditions and served as a reliable observer before conducting formal sessions. In addition, specific training activities were employed, including demonstrating the experimental protocol in informal sessions, to insure that an experimenter would respond appropriately during each assessment condition.

Both graduate and undergraduate students served as primary and reliability observers. Undergraduate students were enrolled in or had completed a course in behavior analysis. Observers received written response definitions and observed each subject with an experienced observer to insure they could identify each response. New observers

collected data in sessions with reliable observers until 90% or better interobserver reliability was obtained on three sessions for each subject.

Each subject's SIB was rated on the MAS by two staff who had worked with the subject extensively (from several months to several years). One supervisor and one direct-care staff were selected for each subject. A total of 24 staff participated as responders for the 34 subjects.

Reliability

Inter-rater reliability for the MAS was computed using two methods: point-to-point agreement for each question, and agreement on the category rankings by each of the respondents. Point-to-point reliability was calculated by dividing the number of agreements on the exact answer for each question over the total number of agreements and disagreements. The range of inter-rater reliability was from 0 to 63% agreement, with a mean of 23.4%. Agreement for the highest ranked categories ranged from 0 to 100%. Seventeen of the thirty-four pairs of raters agreed with respect to the highest ranked category. No pair of responders agreed on the rankings for every category for any subject. There was no special training of raters, and reliability could not be improved due to the nature of the questionnaire.

Interobserver agreement for the time-referenced and behavior-referenced assessments was computed as agreement on occurrence. Reliability percentages for SIB were calculated

by first dividing the observation session into consecutive 10-s intervals. For each interval, the smaller rate was divided by the larger rate. These quotients were then averaged for the session and the total was multiplied by 100. Interobserver reliability data were collected for 24.6% of the total number of time-referenced assessment sessions for all subjects. The mean reliability percentage was 98.6% for the time-referenced assessments. Interobserver reliability was collected for 22% of the total behavior-referenced assessment sessions. The average percentage of interobserver agreement for the behavior-referenced assessment was 95.5%.

Experiment 1

Subjects

Thirty-four subjects (14 males and 20 females), ranging in age from 19 to 45 years, participated in the MAS and the time-referenced assessment comparison.

Experimental Design

The MAS was administered, prior to beginning the time-referenced assessment, to two staff persons familiar with the each subject. All of the subjects were exposed to each of the three different conditions of the time-referenced assessment in a multielement experimental design (Sidman, 1960). The order of presentation of conditions for each subject was random with the restriction that a condition could not be repeated unless each of the 3 conditions had been presented the same number of times.

Results

Figures 2 and 3 show the mean ratings across categories for each subject's MAS, and the mean responses per minute for the time-referenced assessment conditions. The corresponding points for each assessment are connected by vertical lines, indicating that the MAS category and time-referenced assessment condition relate to similar variables controlling the SIB. The time-referenced assessment had no condition corresponding to the tangible category of the MAS; therefore, no point for that category is indicated on any of the graphs. For example, (see Figure 2) Subject 5's MAS ratings are all

high, with means ranging from 3.9 to 4.9. The attention category had the highest mean rating of 4.9. The time-referenced assessment results indicate that the highest mean rate of SIB occurred in the attention condition (.20 responses per minute). The other conditions had the same low rate of SIB, .1 responses per minute.

Fourteen of 34 subjects had the highest mean rating for the MAS in the automatic reinforcement category (sensory stimulation). Twenty-two of the subjects' mean MAS ratings were essentially undifferentiated across conditions (i.e., at least three of the four ratings within 1 point of each other). Comparison of the MAS ratings and the time-referenced assessment reveals that for 8 of the 34 (23.5%), MAS ratings matched the results obtained in the time-referenced assessment (i.e., the highest mean rating for the MAS corresponded to the condition with highest mean rate of SIB for the time-referenced assessment). The subjects for whom results of both assessments matched were: S2, S5, S7, S9, S12, S18, S23, and S25.

Figure 2. Mean ratings for the MAS categories and mean responses per minute for the time-referenced assessment conditions are presented for Subjects 1 - 17. MAS ratings are presented on the left y-axis, with scores ranging from 0 to 6. The mean responses per minute are presented on the right y-axis, with the range individually determined by each subject's rate of responding.

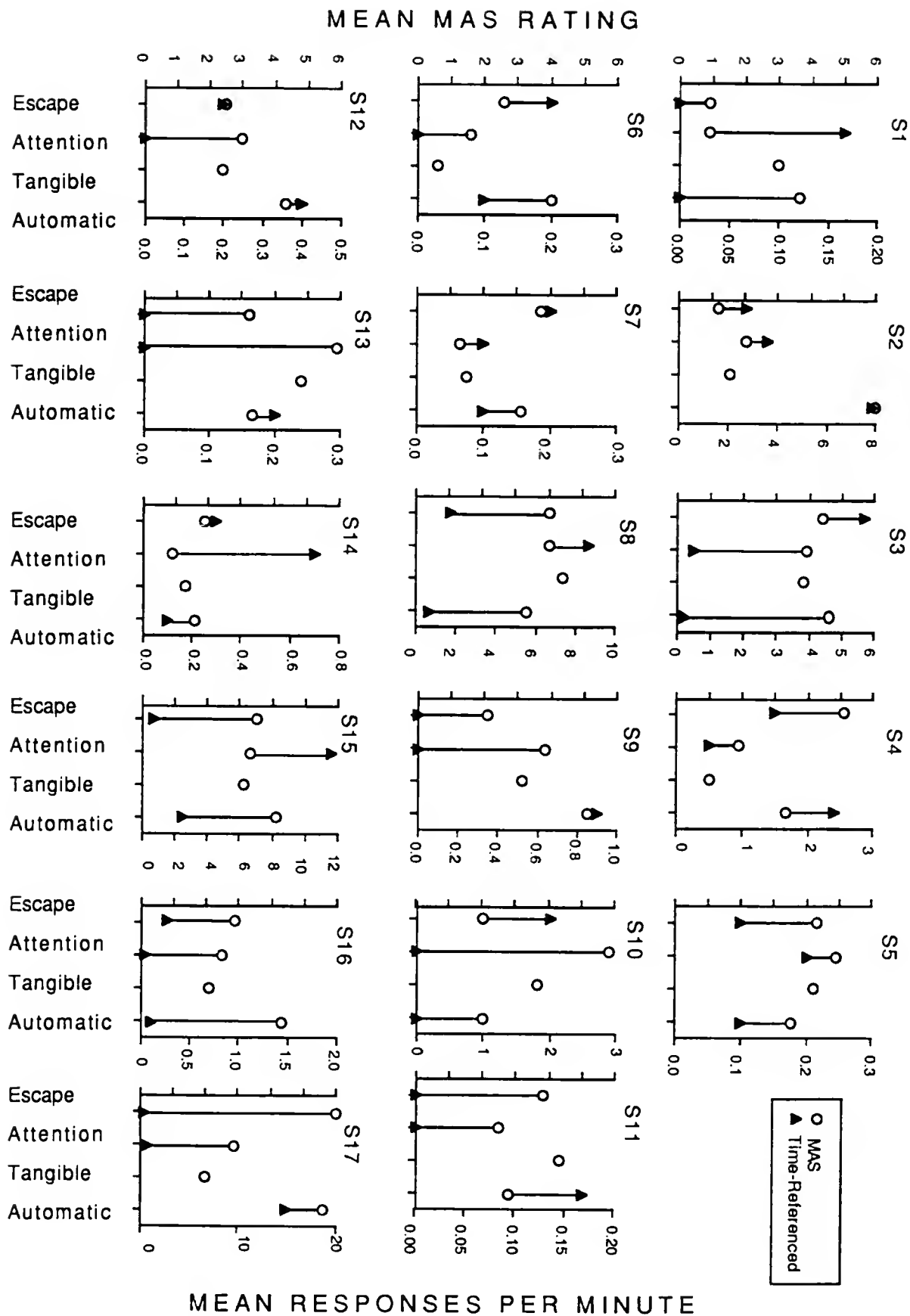
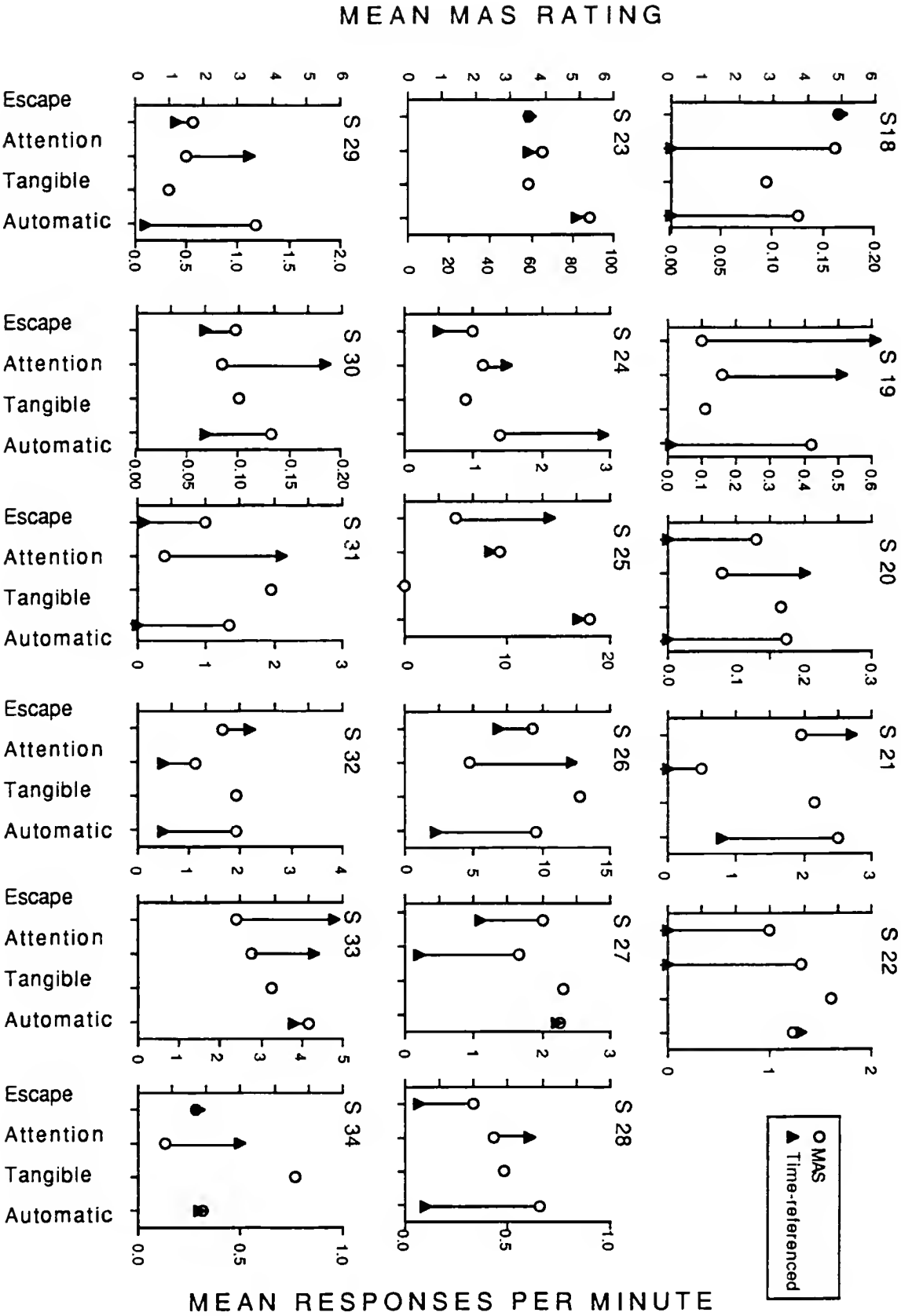


Figure 3. Mean ratings for the MAS categories and mean responses per minute for the time-referenced assessment conditions are presented for Subjects 18 - 34. MAS ratings are presented on the left y-axis, with scores ranging from 0 to 6. The mean responses per minute are presented on the right y-axis, with the range individually determined by each subject's rate of responding.



Spearman rank order correlations based on ranked means for the MAS ratings and the ranked mean rates of SIB for the Demand/Escape, Attention and Sensory/Alone conditions were completed for all 34 subjects. The Tangible condition of the MAS was not included in the correlations because there was no directly corresponding condition in the time-referenced assessment. The range was -0.87 to 1.0, with a mean of 0.09.

Discussion

This experiment was a systematic replication of the Durand and Crimmins (1988) comparison between a time-referenced analogue assessment and an indirect assessment technique (the MAS). The data indicated that the MAS is not an accurate predictor of the results of an analogue assessment; that is, there was poor correspondence between MAS ratings and the rates of SIB during analogue conditions of the direct assessments.

Durand & Crimmins presented the MAS as an alternative method for identifying the variables controlling SIB. They described two major drawbacks of the analogue method of functional analysis: 1) The analogue assessment requires extensive staff training to implement; and 2) The analogue assessment requires an extensive amount of time. Although the MAS might alleviate these problems, several more serious problems arise as a result of its use. First, the MAS also might require that extensively trained staff are chosen to complete the questionnaire. This supposition is based on the

low inter-rater reliability results obtained in the present study and a prior study (Zarcone et al., in press). Second, the results of the MAS for the current study were poor predictors for the results of an analogue assessment. Therefore, the validity of the instrument for identification of controlling variables of SIB is questionable. The poor interrater reliability may have been a factor in the low validity obtained in the current study. It should be noted that these results may not be indicative of the validity of the MAS given better interrater reliability.

An additional point should be made with respect to the Durand & Crimmins study. Rather than selecting subjects for their analogue comparison randomly, they selected 2 subjects with high ratings in a single category for each of the four MAS categories. This selection should have insured that the controlling variables for the SIB would be easily discovered with the analogue assessment; however, of the 8 subjects selected by Durand and Crimmins, differentiation among analogue conditions occurred for only 2 subjects.

Experiment 2

Subjects

Eight subjects (2 males and 6 females) participated in the time-referenced analogue assessment and the behavior-referenced analogue assessment comparison. These 8 subjects were chosen from the 34 subjects who participated in Experiment 1. These subjects were selected because their

high rate or high degree of severity of their SIB targeted them as priority candidates for treatment.

Experimental Design

The time-referenced assessment (from Experiment 1) was completed prior to beginning the behavior-referenced assessment for each subject. For 6 of the 8 subjects, the four experimental conditions were presented in a multielement design. The order of presentation of conditions for each subject was sequenced in a random order with the restriction that one condition could not be repeated until all three conditions had been presented. Three to four sessions occurred each day. The behavior-referenced assessment was in effect until differential responding occurred, or until at least 12 sessions of each condition had been completed. For the remaining 2 subjects, (19 and 32) the conditions of demand, attention, and play were presented in an A-B-C design. The sequence of presentation of conditions differed for these subjects, and each condition was continued until there was some stability in responding from session to session, or for at least eight sessions. Stability was determined based on eyeballing the rates per minute from individual session.

Results

Figure 4 displays the mean responses per minute from the time-referenced and the behavior-referenced assessment conditions for each of the subjects. The mean responses for

each condition are presented in order to allow a concise comparison for both of the assessments. The individual session data for each assessment would have resulted in a comparison of 9 sessions for the time-referenced assessment and for several subjects, 40 to 176 sessions for the behavior-referenced assessment.

Five of the subjects had their highest mean rates of responding for the same conditions in the time-referenced and behavior-referenced assessment. These subjects are indicated by an asterisk above the highest conditions on each graph (see Figure 4). The play condition resulted in the lowest mean rate of SIB for two of the six subjects (16 and 24), for whom a play condition was implemented.

Subject 8 exhibited SIB most frequently during the attention condition of both assessments, low mean rates of responding in the alone and demand conditions of both assessments. The rate of responding during the play session was slightly higher than either the alone or demand conditions.

Subject 15 had the highest mean rate of SIB in the attention conditions of both the time-referenced and behavior-referenced assessments. There were low rates in all other conditions.

Subject 16 had low rates of SIB in all conditions of the time-referenced assessment, with the highest rates occurring in the demand condition. Low rates of SIB occurred

in the attention, alone, and play conditions of the behavior-referenced assessment. The highest mean rate of SIB was in the demand condition, with a difference of four times the rate between that condition and the condition with the next highest mean rate of responding.

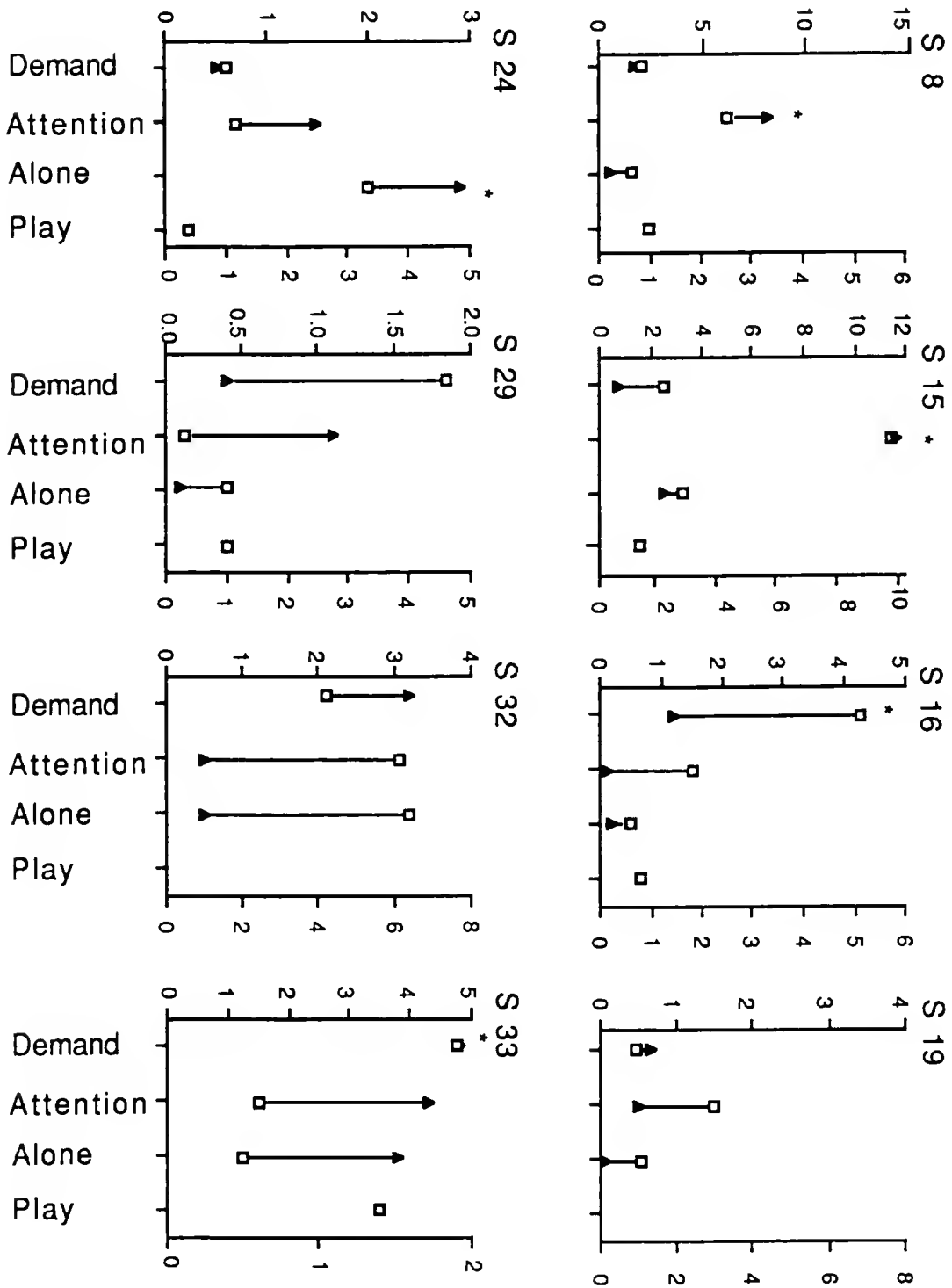
Subject 19 had low mean rates of SIB in all conditions of the time-referenced assessment. The rates for the behavior-referenced assessment were highest in the attention condition with low rates in the demand and alone condition. The play condition was not included in either assessment.

Subject 24 had the highest mean rate of SIB in the alone condition of the time-referenced assessment. The next highest rates occurred in the attention condition and demand conditions, respectively. Results for the time-referenced and behavior-referenced assessments concurred across all conditions. The alone condition resulted in the highest rates in both assessments, and the lowest rate occurred in the control, behavior-referenced condition.

Subject 29 exhibited the highest mean rate of SIB in the attention condition of the behavior-referenced assessment. During the behavior-referenced assessment, the highest mean rate of SIB occurred during the demand condition with the lowest mean rate of SIB occurring in the attention condition.

Figure 4. Mean responses per minute for the time-referenced assessment and the behavior-referenced assessment are presented for each subject. On the left y-axis are the mean responses for the time-referenced assessment with the ranges individually determined by each subject's rate of responding. On the right y-axis are the mean responses for the behavior-referenced assessment with the ranges individually determined by each subject's rate of responding.

Mean Responses per minute Time-Referenced Assessment



Mean Responses per minute Behavior-Referenced Assessment

Subject 32 exhibited the highest mean rates of SIB in the demand condition of the time-referenced assessment and the alone condition of the behavior-referenced assessment. The difference between the attention condition rate and the rate in the alone condition was small.

Subject 33 exhibited the highest rates of SIB in the demand condition of both assessments. There was little difference between the rates for the attention and alone conditions of either the time-referenced or the behavior-referenced assessments. The rate of SIB in the play condition was higher than the rate in either the alone or attention condition of the behavior-referenced assessment.

A comparison between the results of the MAS and the results of the behavior-referenced assessments for these 8 subjects is possible. For Subject 8 (see Figure 2) the MAS rating was highest for the tangible category and the mean rate of responding was highest in the attention category of the behavior-referenced assessment (see Figure 4), indicating a possible match with positive reinforcement as the identified maintaining variable; for Subject 24 the MAS category with the highest mean rating was automatic (see Figure 3), and the behavior-referenced assessment condition with the mean highest rate of responding was alone (see Figure 4), indicating a possible match with automatic reinforcement as the identified maintaining variable. The

MAS results for Subject 32 were regarded as not matching the behavior-referenced assessment because two categories were rated equally high for the MAS (automatic and tangible), while responding in the behavior-referenced assessment was highest in the attention condition. Based on this comparison, the MAS identified the maintaining variable for 2 of the 8 subjects.

Spearman rank order correlations based on the ranked mean responses per minute of the time-referenced assessment and the behavior-referenced assessment for the Escape, Attention, and Alone conditions were completed for the 8 subjects participating in the study. The Play condition of the behavior-referenced assessment was not included because there is no corresponding condition in the time-referenced assessment. The range was -0.5 to 1.0, with a mean of 0.12. In addition, Spearman rank order correlations for the ranked mean MAS ratings and the ranked mean responses per minute of the behavior-referenced assessment for the Demand/Escape, Attention and Sensory/Alone conditions were completed for each of the eight subjects. The range was -1.0 to 1.0, with a mean of -0.06.

Discussion

This experiment compared the results of brief exposure to contingencies arranged in several analogue conditions to the results of more lengthy exposures to the same conditions. The question was whether the time-referenced assessment would

provide results similar to those of the behavior-referenced assessment, thereby providing a more efficient functional analysis of the variables maintaining SIB. Of the 8 subjects for whom both assessments were implemented, the time-referenced assessment condition with the highest mean rate of SIB was the same as the behavior-referenced assessment condition in five cases. That is, the highest mean rates were the same for approximately 60% of the subjects.

Because the subjects participating in this experiment also participated in Experiment 1, the MAS and behavior-referenced assessment results were compared. Such a comparison revealed that the MAS ratings of staff agree with the results of an experimental manipulation of environmental variables for only 2 of the 8 subjects, a 25% agreement. The MAS results do not agree with either time-referenced or behavior-referenced analogue assessment results

General Discussion

The present study compared results obtained from three methods for assessing functional properties of SIB: an interview procedure (the MAS) and two direct methods (time-referenced and behavior-referenced assessments) in which environmental conditions were systematically controlled. A comparison of the MAS and a time-referenced analogue assessment revealed that the MAS results did not match the time-referenced results.

Experiment 1 was a systematic replication of the Durand and Crimmins (1988) study. There were two major differences between the Durand & Crimmins and the present study, and these differences may have been at least partially responsible for the results obtained in the current study. First, the demand condition in the Durand & Crimmins study was "aversive" along the dimension of difficulty, as indicated by the probability of a correct response on the items, and possibly along the dimension of time-out from the task/attention. In the current study, other dimensions of "demand" that might be aversive to subjects were included: effort of response, physical prompting, and demands other than academic tasks. The second difference between the present study and the Durand & Crimmins study was the selection of subjects. For each MAS category, Durand & Crimmins selected two subjects whose MAS ratings were high for that category and that category alone. The current study selected subjects in a quasi-random manner. Subjects were screened as they were referred to an SIB intervention project and were included in the study as the assessment portion of their treatment was completed. Therefore, the current study did not specifically select subjects based on any characteristic criteria of their SIB. The quasi-random selection may have resulted in less distinctive performance during assessment sessions, as the subjects' SIB may have

been less differentiated than others who scored particularly high in an MAS category.

Experiment 2 involved a comparison of two types of functional assessments: brief exposure to experimental conditions in the time-referenced assessment and lengthy exposure to the same conditions in the behavior-referenced assessment. The time-referenced results matched the behavior-referenced assessment results for 5 of the 8 subjects. These results suggest that the accuracy of an abbreviated functional analysis may be limited; correctly identifying the maintaining variable 62.5% of the time is a low success rate. The brevity of a subject's contact with the contingencies, (through few sessions and brief exposure) may affect the comparison of the subject's behavior during sessions and in his/her normal environment. This is a familiar problem of sampling; the larger the sample, the more representative it will be of the total population. In the case of a single subject's behavior, more frequent measurement, in more conditions that approximate all possible instances, will yield more accurate data. The time-referenced assessment of the current study, and the brief analogue assessment of Durand and Crimmins (1988), Steege et al. (1989) (1990), Sturmey et al. (1988), and Wacker et al. (1990) studies, maybe too brief to provide a consistently representative picture of a subject's behavior. The time-referenced assessment described in the present study, or some

variation thereof, may be useful as a "screening" technique; however, it may not be a substitute for a more complete functional analysis.

As previously discussed, a possible criticism of the analogue assessment is that the subject is trained to exhibit the inappropriate behavior by the contingencies manipulated during the assessment. The training would be evident in the data in the form of an acquisition curve. The subjects in the current study did not increase their rate of SIB over the course of the analogue sessions. Instead, a high rate of SIB occurred in one of the conditions and was maintained, the rate of SIB in the other conditions decreases as discriminated responding developed. Therefore, there was no evidence that subjects were trained to exhibit SIB under the control of new variables.

The criticism that the complexity of the analogue method of functional analyses prohibits its use in most settings has led to the use of less accurate methods for identifying the function of SIB, or to applying contingencies without regard to the maintaining variables. A need exists for a more easily utilized method of determining the controlling variables for inappropriate behaviors. It must be pointed out that the treatment of SIB is very complex and made more so if unnecessary or ineffective procedures are implemented. In addition, choosing a treatment on the basis of the erroneous identification of a maintaining variable may be

harmful to the subject. For example, if a subject's SIB was determined to be sensitive to negative reinforcement (avoidance/escape), and the assessment inaccurately identified the variable as positive reinforcement (attention), a time out procedure might be implemented. The use of this procedure would constitute negative reinforcement and the subject's SIB would increase in rate or severity as a result of treatment.

There were two possible limitations of the current studies. First, the poor interrater reliability for the MAS in Experiment 1 may have biased the validity of the MAS-Time-referenced assessment comparison. The interrater reliability could not be improved without significant complications. Durand & Crimmins (1988) may have had better interrater reliability because their responders were better observers of behavior, or because their responders were able to complete the questionnaire while observing the subject, perhaps arranging conditions to test for some of the answers. If either of these possibilities was responsible for the difference, the MAS no longer has the advantage of being less cumbersome than analogue studies and not requiring trained observers.

A limitation of Experiment 2 is that eight subjects is too few to make a conclusive statement regarding the validity of the time-referenced assessment as compared to the behavior-referenced assessment. Completing the comparison

with more subjects may alter proportion of matches between the assessments.

Another type of assessment that eliminates some of the criticisms associated with the functional analysis, and that may also eliminate the problems inherent to the indirect assessments, is the naturalistic observation method. The advantages of naturalistic observation include ease of use because there is no need to control the environment, the efficiency with respect to time and personnel involved, and the minimal disruption of the subjects' lives. Many problems might be resolved if a more standardized method of observation were developed that maintained these advantages. A naturalistic observation method may be an important adjunct to the analogue method of functional analysis, in that more idiosyncratic variables acting on the behavior of interest might be identified. Furthermore, identification of these variables during the assessment would make it easier to develop an effective treatment, and would facilitate the generalization of treatment effects to the natural environment.

Further research investigating the effects of more extensive exposure during time-referenced assessments, as well as combinations of naturalistic observations and analogue assessments, is needed.

REFERENCES

- Bijou, S.W., Peterson, R.F., & Ault, M.H. (1968). A method to integrate descriptive and experimental field studies at the level of data and empirical concepts. Journal of Applied Behavior Analysis, 1, 175-191.
- Carr, E.G. (1977). The motivation of self-injurious behavior: A review of some hypotheses. Psychological Bulletin, 84, 800-816.
- Carr, E.G., Newsom, C.D., & Binkoff, J. (1980). Escape as a factor in the aggressive behavior of two retarded children. Journal of Applied Behavior Analysis, 13, 101-117.
- Chamberlain, P. & Reid, J.B. (1987). Parent observation and report of child symptoms. Behavioral Assessment, 9, 97-109.
- Demchak, M.A. & Halle, J.W. (1985). Motivational assessment: A potential means of enhancing treatment success of self-injurious individuals. Education and Training of the Mentally Retarded, 20, 25-38.
- Durand, V.M., & Crimmins, D.B. (1988). Identifying the variables maintaining self-injurious behavior. Journal of Autism and Developmental Disorders, 18, 99-117.
- Iwata, B.A., Dorsey, M.F., Slifer, K.J., Bauman, K.E., & Richman, G.S. (1982). Toward a functional analysis of self-injury. Analysis and Intervention in Developmental Disabilities, 2, 3-20.

- Iwata, B.A., Vollmer, T.R., & Zarcone, J. R. (1990). The experimental (functional) analysis of behavior disorders: Methodology, applications, and limitations. In A.C. Repp & N.N. Singh (Eds.), Perspectives on the use of nonaversive and aversive interventions for persons with developmental disabilities (pp. 301-330). Sycamore Il: Sycamore Publishing Co.
- Johnson, W.L., & Baumeister, A. (1978). Self-injurious behavior: A review and analysis of methodological details of published studies. Behavior Modification, 2, 465-484.
- Johnson, S.M., & Bolstad, O.D. (1973). Methodological issues in naturalistic observation: Some problems and solutions for field research. In L.A. Hamerlynck, L.C. Handy, & E.J. Mash (Eds.), Behavior Change: Methodology, Concepts, and Practice (pp. 7-67). Champaign, Ill.: Research Press.
- Mace, F.C., & Knight, D. (1986). Functional analysis and treatment of severe pica. Journal of Applied Behavior Analysis, 19, 411-416.
- Repp, A.C., Felce, D., & Barton, L.E. (1988). Basing the treatments of stereotypic and self-injurious behavior on hypotheses of their causes. Journal of Applied Behavior Analysis, 21, 281-290.
- Sidman, M. (1960). Tactics of scientific research. New york: Basic Books.
- Steege, M.V., Wacker, D.P., Berg, W.K., Cigrand K.C., & Cooper, L.J. (1989). The use of behavioral assessment to prescribe and evaluate treatments for severely handicapped children. Journal of Applied Behavior Analysis, 22, 23-33.
- Steege, M.V., Wacker, D.P., Cigrand, K.C., Berg, W.K., Novak, C.G., Reimers, T.M., Sasso, G.M., & DeRaad, A. (1990). Use of negative reinforcement in the treatment of self-injurious behavior. Journal of Applied Behavior Analysis, 23, 459-468.
- Sturmey, P., Carlsen, A., Crisp, A.G., & Newton, J.T. (1988). A functional analysis of multiple aberrant responses: A refinement and extension of Iwata, et al.'s methodology. Journal of Mental Deficiency Research, 32, 31-46.

- Touchette, P.E., MacDonald, R.F., & Langer, S.N. (1985). A scatter plot for identifying stimulus control of problem behavior. Journal of Applied Behavior Analysis, 8, 343-351.
- Wacker, D.P., Steege, M.W., Northup, J., Sasso, G., Berg, W., Reimers, T., Cooper, L., Cigrand, K., & Donn, L. (1990). A component analysis of functional communication training across three topographies of severe behavior problems. Journal of Applied Behavior Analysis, 23, 417-430.
- Zarcone, J.R., Rodgers, T.A., Iwata, B.A., Rourke, D.A., & Dorsey, M.F. (in press). Reliability analysis of the Motivation Assessment Scale: A failure to replicate. Research in Developmental Disabilities.

BIOGRAPHICAL SKETCH

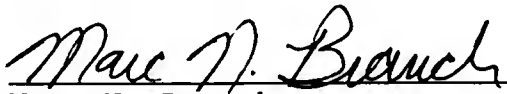
I was born on July 27, 1957, in Madison, Wisconsin. My father was a technician in the Air Force, and my family lived in a variety of places including Japan, Montana, Colorado, Florida and New York. In 1975 I graduated from Peru High School in Peru, New York. I received my Bachelor of Science degree in 1979 from Framingham State College in Massachusetts. I attended the University of Central Florida from 1979 to 1981, and received my Master of Science degree in clinical psychology. I moved to Gainesville, Florida, in 1981 and held a variety of professional positions including psychologist at the Tacachale institution. I returned to school in 1984 in order to pursue a Ph.D. degree. While attending school I was employed as a psychologist at Tacachale, for two years. I was a teaching assistant for three years and have been a research assistant on the Self-Injurious Behavior Intervention Project for three years. In 1988 I married Michael Stoutimore.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the Degree of Doctor of Philosophy.



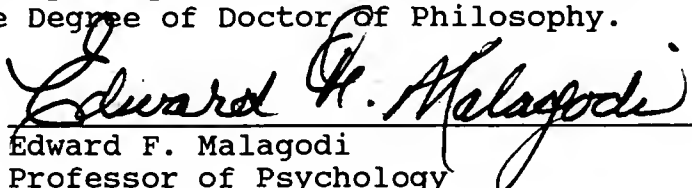
Brian A. Iwata, Chairperson
Professor of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the Degree of Doctor of Philosophy.



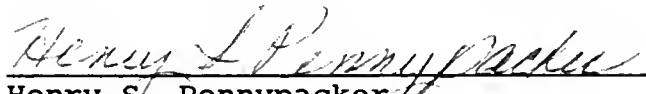
Marc N. Branch
Professor of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the Degree of Doctor of Philosophy.



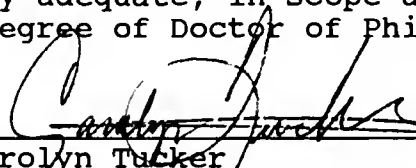
Edward F. Malagodi
Professor of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the Degree of Doctor of Philosophy.




Henry S. Pennypacker
Professor of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the Degree of Doctor of Philosophy.



Carolyn Tucker
Professor of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the Degree of Doctor of Philosophy.



William D. Wolking
Professor of Special Education

This dissertation was submitted to the Graduate Faculty of the Department of Psychology in the College of Liberal Arts and Sciences and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August, 1991

Dean, Graduate School

